Apr 15th, 12:00 AM

The near future of science and technology resources in Japan

Kakugyo S. Chiku
Kanazawa Institute of Technology

THE NEAR FUTURE OF SCIENCE AND TECHNOLOGY RESOURCES IN JAPAN

Kakugyo S. Chiku

Director of Subject Librarian Division
Kanazawa Institute of Technology Library Center, Kanazawa, Japan

Today, Japan is one of the world's most advanced countries in the field of science and technology. This owes much to the fact that, since World War II, not only the universities, research institutes and business corporations, but the state itself, have been eager to collect information on science and technology. The first organ representing the state's effort is the Science and Technology Resources Section of the National Diet Library (NDL). The second is the Japan Information Center of Science and Technology, and the third is the System for Academic Information Center, which was established only recently. I would like to give today an outline of these three institutions.

I. Science and Technology Resources in the National Diet Library

The collection of National Diet Library was started in 1952, when it began collecting the U.S. PB reports. NDL started to collect resources concerning atomic energy in 1953, and Atomic Energy Commission Reports (DOE) in 1954. The collection of foreign journals amounted to 250 titles, thus, the basis for the collection of science and technology resources was established.

In 1959, a legislative act that science and technology resources should be concentrated at NDL was approved in the Diet. This suggests that the collection of science and technology resources by NDL is based on a national policy of developing science and technology in Japan. The collection of foreign journals continued to increase and reached the level of 3,000 titles in 1960. In 1961, the three-year project of science and technology resources began, according to the act of 1959. The aims of the project were, firstly, to collect 10,000 titles of science and technology journals from abroad, secondly, to have a systematic collection of technical reports, and thirdly, to complete an index of journal articles, technical reports and patents.

NDL made efforts to collect resources based on the project. However, from 1965 the budget allocated for journals did not keep up with soaring prices. Furthermore, as the price of journals increased considerably after the oil crisis of 1972, the purchase of many journals was cut off in the following years. And so, in 1978 the collection of journals was reduced to the level of 6,000 titles.

In 1979, as a result of an appeal to the Diet by NDL, the budget for the purchase of science and technology resources was approved and a substantial increase was granted.

Supported by this approval, NDL made another plan for the collection of journals from abroad, setting its goal at the level of 45,000 titles, which was equal to the collection of British Library Lending Division (BLLD). The collection plan had three major policies:

1) The collection was to cover all areas of science and technology;
2) It was to choose core journals thoroughly and preferentially;
3) It was to collect special kinds of resources which were difficult for other institutions to obtain, such as the bulletins of universities and various research institutes, and also journals of special districts in special languages.

In 1979 NDL collected 4,676 titles, focusing on English-language versions of Soviet journals, and in 1980 another 5,693, focusing on the field of clinical medicine. An additional 2,823 titles were collected in 1981, giving priority to journals of the Communist World, such as Trudy, and Sbornik. Sometimes it was impossible to obtain such journals in Japan. Furthermore, it takes a very long time to obtain journals from abroad, in some cases, requiring a few years. The NDL collection of foreign science and technology journals amounted to 10,446 titles in December 1981, and the project of collecting 45,000 titles continues to be carried out.

However, as the Japanese economy has entered a period of stable growth and the state budget is to be formed with a zero or minus ceiling, the budget for journals has not increased very much. In 1984, the NDL collection was approaching 13,000 titles of foreign science and technology journals and 5,600 titles of Japanese ones. And the collection of core journals has reached a considerable level.

The NDL collection at the present stage and its range concerning abstracts, journals, and index journals are shown in Table 1.

Table 1: NDL coverage for the entries of major abstract and index journals.

The entries of major abstract & index journals		NDL coverage	
Chemical Abstracts	17,863 titles	11,000 titles	62%
Science Abstracts (INSPEC)	2,432	2,429	99%
Engineering Index	1,974	1,780	90%
Science Citation Index	3,300	3,250	98%
INIS Atomindex	2,206	2,100	95%
International Aerospace Abstracts	1,100	1,096	99%
Metals Abstracts	705	680	96%
Index Medicus	2,561	2,200	86%
Excerpta Medica	4,415	3,300	75%

"Science and Technology Information Service", No. 59, Feb. 1982, National Diet Library.

Table 2 and the following describe the collection of "Gray literature" in NDL.

Table 2: Collection of "Gray literature" in NDL

1. Technical Reports

U. S. A.	AD, PB Reports	National Technical Information Service (NTIS)	1945-	660,809 items
U. S. A.	OSRD Reports	Office of Scientific Research and Development	1941-45	13,855
U. S. A.	TOM Reports	Technical Oil Mission	1946-50	16,509
U. S. A.	DOE Reports (AEC, ERDA)	U.S. Dept. of Energy	1947-	293,395
U. S. A.	DOCKET Reports	U.S. Nuclear Regulatory Commission	1967-	194,285
U. K.	AEA Reports and others	Atomic Energy Authority and others	1951-	8,411
CANADA	AECL Reports	Atomic Energy of Canada, Ltd.	1961-	4,771
FRANCE	CEA Reports and others	Commissariat á l'Energie Atomique and others	1948-	6,462
B. R. D.	HMI Reports and others	Hann-Meitner-Institut für Kernforschung and others	1974-	2,116
ITALY	CNEN-RT Reports	Comitato Nationale per l'Energia Nuclear	1960-	5,786
POLAND, INDIA, and other countries	PAN, BARC, A/AC 82 Reports and others			13,914
INIS	INIS Reports	International Nuclear Information System	1979-	75,197
U. S. A.	NASA Reports (NACA)	National Aeronautic and Space Administration	1959- (1915-58)	80,000
U. K.	ARC Reports	Aeronautical Research Council, Ministry of Aviation	1929-80	4,350
FRANCE	ONERA Reports	Office National d'Etudes et Recherches Aerospatiales	1947-	500
FRANCE	PSTMA Reports	Ministére de l'Air	1930-62	500
U. S. A.	RAND Reports	RAND Corporation	1964-	10,105
U. S. A.	MIT Reports	Massachusetts Institute of Technology	1951-63	825

2. Dissertations

U.S.Dissertations	University Microfilms International	1979-	75,197 items
U.K.Dissertations	University of Oxford, University of London and others	1967-52 1976-77	905
B.R.D.Dissertations	Universität Karlsruhe, Johan Wolfgang Goethe Universitat and others	1954-	20,052
French Dissertations	Université du Paris, Inst. National Polytechnique de Toulouse and others	1956-60 1967-70 1975-	7,505
Dutch Dissertations	Rijksuniversiteit te Leiden, Technische Hogescool Delft	1960-	1,457
Swiss Dissertations	Eidgenössische Technische Hochschule, Zürich, UniversitätBasel	1960-	1,285
Swedish Dissertations	Chalmers Tekniska Högskola and others	1970-	504
Finnish Dissertations	Finland Inst. of Technology	1961-	28

3. Institutional Papers

U. S. A.	AIAA Papers	American Inst. of Aeronautics and Astronautics	1963-	29,727 items
U. S. A.	ASME Papers	American Society of Mechanical Engineers	1958-63 1964-	23,228
U. S. A.	SAE Papers	Society of Automotive Engineers	1959-60 1963-	13,732
U. S. A.	SME Papers (ASTME)	Society of Manufacturing Engineers	1967-72 1982-	2,207
U. S. A.	IEEE Papers	Inst. of Electrical and Electronics Engineers	1965-68	1,486

4. Standards

JAPAN	JIS Standards	Japanese Standards Association
JAPAN	JIS Standards (English version)	Japanese Standards Association
JAPAN	JAS Standards	Japan Agricultural Standards Association
JAPAN	JEC Standards	Inst. of Electrical Engineers of Japan
U. S. A.	FS Standards	General Services Administration
U. K.	BS Standards	British Standards Institution
CANADA	CGSB Standards	Canadian Government Specifications Board
CANADA	CAN Standards	Standards Council of Canada
B. R. D.	DIN Standards	Deutscher Normenausschuss
	ISO Standards	International Organization for Standardization
	IEC Standards	International Electrotechnical Commission

5. Foreign Patent Details

U. S. A.	U.S. Patents	U.S. Patent Office	1952-65	206,000 items
	PCT Patents	World Intellectual Property Organization	1978-	15,618
	EPO Patents	European Patent Office	1982-	12,357

6. Fact Data

U. K.	ESDU	Engineering Sciences Data Unit Ltd.		
U. S A.	Sadtler Standard & Commercial Spectra	Sadtler Research Laboratories		
U. S. A.	X-ray powder data file	American Society for Testing and Materials		28,500 items
JAPAN	IRDC-Cards	Infra-Red Data Committee	1960-77	19,200

"Science and Technology Information Service", No. 67, Feb. 1984, No. 70, Nov. 1984, National Diet Library.

The collection of science and technology resources in NDL is the largest in Japan, and is fully used. There is no doubt that the NDL collection and its service has supported the development of science and technology in Japan. NDL intends to become equal to BLLD in its collection of resources, and also to provide a more efficient service by processing the collection to data base.

II. The Japan Information Center for Science and Technology (JICST)

It seems that JICST has come to be known widely in Europe recently, so here, I would simply like to introduce the future plans of JICST.

JICST was established as a special non-profit organization under a legislative act in 1957. JICST is under the executive control of the Science and Technology Agency, Prime Minister's Office. In 1984, the JICST collection included 6,810 titles of foreign journals on science and technology, 4,050 titles of Japanese journals, 12,500 technical reports, and 50,000 items of patent details. Japanese-language abstracts of these resources are published monthly or semi-monthly as the "Current Bibliography on Science and Technology (CBST)", in which the resources are divided into 12 files. The bibliography and abstracts are processed into a data base, and, along with fifteen data bases such as CA SEARCH, MEDLINE, BIOSIS, INSPEC, make up the JICST On-Line Information System (JOIS) - II. This system provides a nation-wide retrieval service.

Japan is often criticized as being a black hole of information. This means that Japan is eager to acquire information from abroad, but not to unveil information to the foreign community. The truth is, however, that Japan does expose almost all information to the public. The problem is that most of this information is in the Japanese language. Actually, 90% of the technological achievements reported in Japan are written in Japanese. Japanese is not a widely known language in foreign countries, therefore, any endeavor to acquire technical information from Japan is, at once, confronted with this language barrier.

In order to solve this problem, JICST started in 1985 to translate into English JICST Files (on Science and Technology - corresponding to CBST, on Current Science and Technology Research in Japan, and on Medical Science in Japan), and at the same time, started the on-line retrieval service on Japanese data base in West Germany and South Korea. JICST also intends to develop a so-called fact data base, such as thermal physical properties, a mass spectrum data base which is on the way at present, a picture data base of chemical compound structures and crystal structures, and a full text data base. In 1988, JOIS-III is expected to be completed equipped with the retrieval functions of these fact data bases with which it is expected to provide an on-line retrieval service to overseas users.

III. The System for Academic Information Center Project

In 1980, the Academic Council of Japan submitted a report entitled "Future Academic Information Systems of Japan" in response to the enquiry from the Minister of Education. In order to maintain and develop the high-level of Japan's scientific studies, the report pointed out that a system should be urgently established for the dissemination and distribution of abundant and specialized scientific information equally and promptly to many scholars across the nation. In May 1980, based on the report, the Ministry of Education started an investigation into the establishment of "the System for Academic Information Center" and planned the construction of a nationwide on-line network system in university libraries. When the investigation ended in 1983, the Ministry of Education decided to set up the center of the system in the Center for Bibliographic Information (CBI) at the University of Tokyo.

By that time, CBI had finished compiling the data base, "the Union List of Scientific Periodicals", which listed all magazines stored in university libraries throughout Japan. (Of its collection, Western-language magazines numbered approximately 91,551, of which 53,000 were science and technology magazines; Japanese-language magazines accounted for 44,605). It planned to allow university libraries to retrieve the magazines through its data-base service. At the same time, it also undertook the same kind of project, i.e. "shared cataloging" and the data retrieval, in connection with all the data these libraries possessed. Furthermore, it planned to use the computerized system to perform a large part of library business, such as selection, acquisition, and card-cataloging of books in member libraries. For this purpose, CBI added LC-MARC, UK-MARC and JAPAN-MARC to its data base as reference files. This system is an integrated academic information system that covers magazines and books in humanities, social science, natural science and technology.

At present, a fully computerized university library is still rare in Japan, with the exception of our Kanazawa Institute of Technology Library Center. However, libraries where computers perform a part of the tasks are rapidly increasing. It is anticipated that most major university libraries will introduce computer systems in some fields within 10 years from now. The wide use of the CBI system will promote the trend, and this is one of the objects CBI is trying to achieve. As N-1 Network, the network system of computer centers in major national universities, was already working in 1983, it was decided that the CBI system should use this N-1 Network and Expanded N-1 Protocol. The latter was specially developed for this project and possessed a function for processing Chinese characters and figures. The local systems in university libraries are equipped with Expanded N-1 Protocol and connected with the central system of CBI via DDX, Packet Exchange Network of NTT (Nippon Telephone & Telegraph Corporation). Figure 1 illustrates this system. It is planned that the system will be developed into a kind of VAN (Value Added Network) in the future: the central system is equipped with X 29 Protocol

of C-CITT, and connected with the local systems, into which PAD (Packet Assembly and Disassembly) is loaded, via DDX. This indicates a possibility that the CBI system can be hooked up to various VAN information systems, including foreign ones.

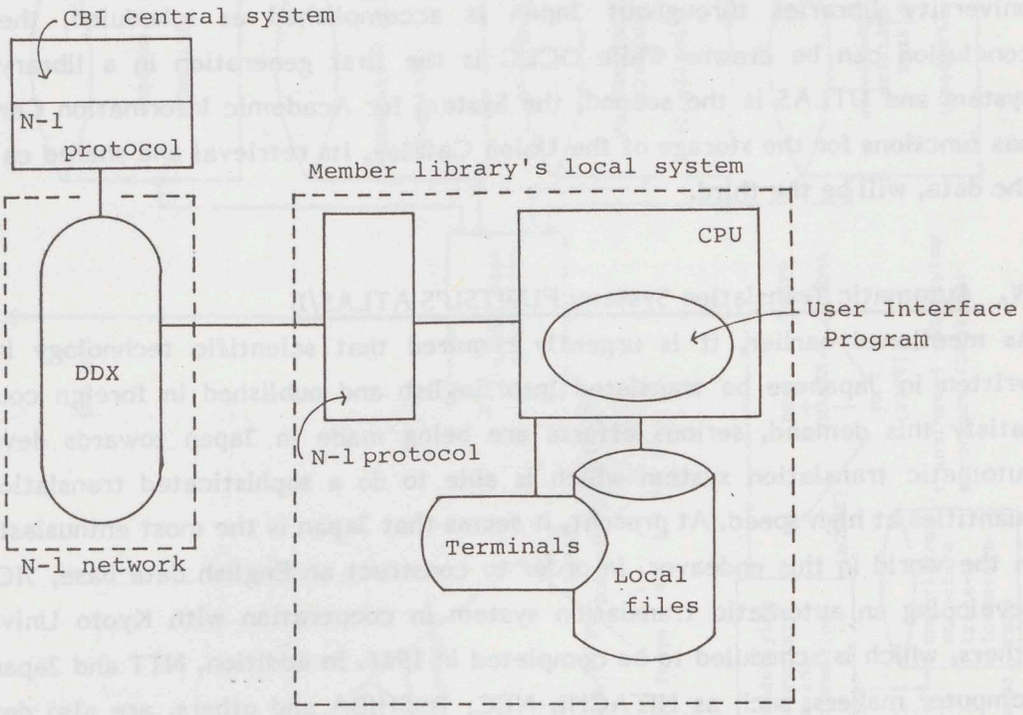


Fig. 1. CBI System flow chart: N-1 System

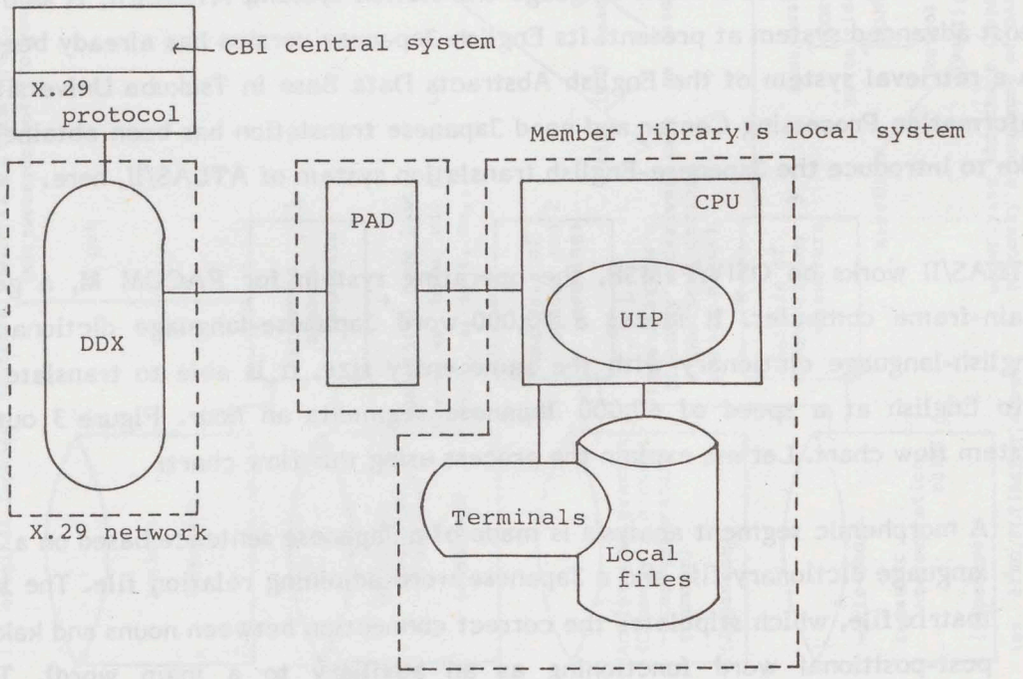


Fig. 2. CBI System flow chart: PAD System (VAN)

In April 1985, CBI was fully connected with the Tokyo Institute of Technology Library. In addition, it is planned that CBI will be linked to national universities such as Kyoto University, Kyushu University and Hokkaido University, and private ones such as Keio University, Nanzan University and Kansai University in 1985. If the VAN system in university libraries throughout Japan is accomplished as scheduled, the following conclusion can be drawn: While OCLC is the first generation in a library computer system and UTLAS is the second, the System for Academic Information Center, which has functions for the storage of the Union Catalog, its retrieval and shared cataloging of the data, will be the third.

IV. Automatic Translation System: FUJITSU'S ATLAS/II

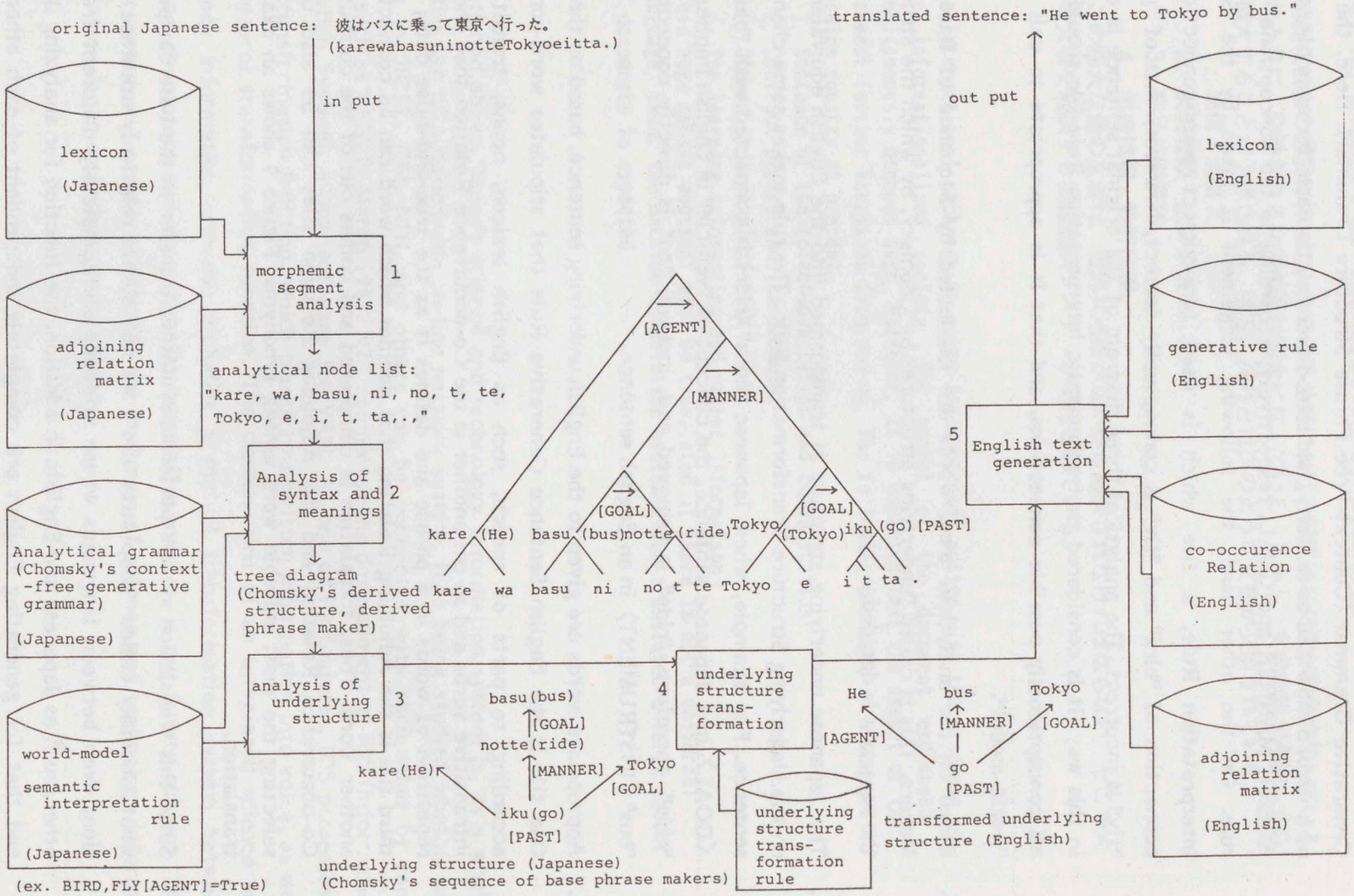
As mentioned earlier, it is urgently required that scientific technology information written in Japanese be translated into English and published in foreign countries. To satisfy this demand, serious efforts are being made in Japan towards developing an automatic translation system which is able to do a sophisticated translation in large quantities at high speed. At present, it seems that Japan is the most enthusiastic country in the world in this endeavor. In order to construct an English data base, JICST is also developing an automatic translation system in cooperation with Kyoto University and others, which is scheduled to be completed in 1986. In addition, NTT and Japan's leading computer makers, such as HITACHI, NEC, TOSHIBA and others, are also developing a translation system.

Among them, FUJITSU's multi-language translation system, ATLAS/II, is said to be the most advanced system at present. Its English-Japanese version has already been included in a retrieval system of the English Abstracts Data Base in Tsukuba University Science Information Processing Center and good Japanese translation has been obtained. I would like to introduce the Japanese-English translation system of ATLAS/II, here.

ATLAS/II works on OSIV/F4MSP, the operating system for FACOM M, a general-use main-frame computer. It stores a 50,000-word Japanese-language dictionary and an English-language dictionary with the same entry size. It is able to translate Japanese into English at a speed of 60,000 Japanese segments an hour. Figure 3 outlines the system flow chart. Let me explain the process using this flow chart:

1. A morphemic segment analysis is made of a Japanese sentence based on a Japanese-language dictionary file and a Japanese word adjoining relation file. The latter is a matrix file, which stipulates the correct connection between nouns and kaku-joshi (a post-positional word functioning as an auxiliary to a main word). Then, the analytical node list of the whole sentence is formed.

Fig 3. FUJITSU's Automatic Translation System: ATRAS/II System Flow



2. The analytical node list is analyzed from the standpoint of connection, grammatical relation and conceptional relation of words, by reference to Analytical Japanese Grammar and Semantic Interpretation Rule based on N. Chomsky's Transformational Generative Grammar (Context-Free Phrase Structure Grammar). Here, the subject of a verb, a means of an activity, and a modifyer relation are, for example, shown as AGENT, MANNER and GOAL, respectively, according to the conceptional relation rule. On the other hand, the following judgement is made by the Semantic Interpretation Rule, a file which is used for correct word-connection: The connection of "bird" and "fly" is conceptually correct, however, that of "pig" and "fly" is incorrect. The analytical three diagram of the original sentence is obtained in this way. It is considered that a Semantic Interpretation Rule file should include all conceptionally possible connections, and that it is appropriate to call the file "world-model".

3. Based on the analytical tree, words are connected by Sequence of Base Phrase Makers. This is called "underlying structure" by Chomsky, by which the structure of the sentence is decided.

4. The Japanese underlying structure is transformed into the English equivalent based on the Underlying Structure Transformation Rule. The rule makes some changes in a sentence. For example, when Japanese "noru" (ride) is connected with "basu" (bus) 〈GOAL〉 in a Japanese sentence, the concept of "vehicle" in English implies that of "ride", therefore, "ride" is removed and, instead, "go" is directly connected with "bus" 〈INSTRUMENT〉 in an English sentence.

5. Appropriate words are given to the English underlying sentence, based on three files: One file is the English Sentence Generative Rule that stipulates word-connection according to parts of speech, such as English lexicon, nouns, transitive and intransitive verbs and so on: another is the Co-occurrence Relation that checks the repetition of words in a phrase and defines it as the true-and-false data; and the third file is the Adjoining Relation that decides which word can be connected with another word. Thus, translation is completed and comes out of the computer. The Co-occurrence Relation and the Adjoining Relation files can be also used for selecting the most suitable word among synonyms. Figure 4 shows an example of translation.

Since this translation system performs syntactic transfer in the base component of what Chomsky called "deep structure", we can obtain relatively accurate translation, even between languages whose syntaxes are completely different from each other, such as Japanese and English. In addition, the function for analyzing Japanese and that for generating English are completely independent of each other in this computer. Therefore, it is easy to develop a multi-language translation system by adding a function for generating other languages.

Fig. 4. An example of translation by ATLAS/II

情報化社会の進展に伴って，あらゆる分野における国際的な交流が活発になっている．それと共に，交換される文書の量は膨大になっている．しかも，文書の量が今後もさらに増大することが予想される．現在，海外とは次のような文書が交換されている．

1．通信・連絡文
2．契約書類・法律文書
3．商品技術資料
4．学術情報・特許情報

この膨大な量の文書は，世界共通語がない限り翻訳されなければならない．そこで大量の文書を翻訳する機械翻訳システムが期待されている．

An international exchange in every field becomes active with the progress of the informationalized society. The amount of the document exchanged simultaneously becomes huge. Moreover, it is expected that the amount of the document further increases from now on. The following document is exchanged to foreign country at present.

1. Sentence of the telecommunication and communication
2. Contract papers and law document
3. Commodity technology data
4. The learning information and patent information

This huge amount document must be translated if there is not a universal language. Then, the machine translation system which translates a large amount of documents is expected.

As mentioned above, Japan's science and technology information system is, as a whole, facing the age of a nationwide on-line network system. It is being studied whether the systems in JICST, CBI, NDL and others can be connected with each other by some means, such as VAN. In March 1985, NTT completed the development of IDS (Information Distribution System), a new software for VAN. As one of its features, IDS possesses a media transformation function, which transforms written information into voice, as well as a protocol transformation function. The construction of an integrated science and technology information network, including a mutual transformation system between voice and characters and an automatic translation system, is our future subject. This network system will greatly accelerate the exchange of science and technology information between Japan and other countries.